A Note on the Computation of the Bit Error Rate for Binary Block Codes*

Michele Elia
Dipartimento di Elettronica
Politecnico di Torino
Corso Duca degli Abruzzi 24
I-10129 Torino, Italy

Submitted by I. F. Blake

ABSTRACT

The bit error rate of binary codes on the binary symmetric channel with raw error probability p is a performance index of primary importance in many applications. It is the probability that decoded information symbols are in error after complete decoding. In this paper the class of linear block codes having a transitive group of automorphisms is considered. It is shown that a linear homogeneous differential transformation of the weight enumerator enables us to obtain a closed expression for the bit error rate in terms of p. The application of the method is general enough to allow the computation of bit error rates for all cyclic codes.

I. INTRODUCTION

Linearity pervades nearly all the manifold aspects of coding theory. The most remarkable classes of error control codes are defined as subspaces of finite vector spaces over a Galois field. The dual code $\mathscr{C}^{\perp}$ of a given linear code $\mathscr{C}$ is defined as the subspace of vectors orthogonal to all vectors of $\mathscr{C}$. The distributions of vector Hamming weights of paired dual codes are connected by a linear transformation. Here linearity will be applied to exploit properties of another important parameter describing code performance, the symbol error probability.

*This paper was presented at IEEE International Symposium on Information Theory, Ann Arbor, Michigan, 4–9 October 1986.

52 Vanderbilt Ave., New York, NY 10017 0024-3795/88/$3.50

Let V_n be the n-dimensional vector space over the Galois field GF(2). Let $\mathrm{wt}(U)$ denote the Hamming weight of the vector $U \in V_n$, i.e. the number of nonzero components of U. The Hamming distance $d(U,V)$ between U and V is defined as the number of components in which the two vectors differ. Equivalently $d(U,V) = \mathrm{wt}(U+V)$. A binary linear code (n,k,d) is a k-dimensional vector subspace of the vector space V_n with minimum Hamming distance between vectors equal d; thus the code can correct $t = \lfloor (d-1)/2 \rfloor$ errors. The weight enumerator is defined as the polynomial

$$W(X,Y) = \sum_{i=0}^{n} A_i X^i Y^{n-i},$$

where A_i is the number of vectors of weight i.

Let $Z[[X,Y]]$ be the algebra of polynomials in X and Y, and let $H_n[[X,Y]]$ be the $(n+1)$-dimensional subspace of homogeneous polynomials $q(X,Y)$ of degree n. A linear transformation on X and Y induces a linear transformation in $H_n[[X,Y]]$. The transformation, denoted M_c and defined as

$$M_c[q(X,Y)] = q(X+Y, X-Y),$$

was introduced by MacWilliams [3] to show the connection between weight enumerators of dual codes. In fact we have the relation

$$W^{\perp}(X,Y) = \frac{1}{2^k} M_c[W(X,Y)].$$

The bit error probability, also called bit error rate and henceforth shortened to BER, is the average probability that an information bit is in error after complete decoding. Complete decoding means that every received word is assigned to a definite codeword. In the case of linear codes of block length n complete decoding is realized through a coset decomposition of V_n with respect to the code. Such a decomposition is usually called a standard array, and different aims may be accomplished by giving different rules for selecting the coset leaders.

Slepian [1] has shown that coset leaders of minimum weight yield a maximum likelihood decoding, thus minimizing the probability that the decoded codeword is not the transmitted one, i.e., the word error probability P_w is a small as possible. In [8] it is shown that maximum likelihood decoding, hereafter abbreviated MLD, does not necessarily minimize the BER. In fact,

for many codes a smaller BER is obtained by taking each coset leader to be either the unique element of minimum weight not greater than t, or the unique element with all zeros in information positions. We shall call this last rule *unique coset leader decoding*, and hereafter it will be abbreviated UCLD.

Sporadic expressions for the BER in terms of p, the raw error probability of the binary symmetric channel, have been derived [2, 5, 6, 7]; in [8] the performances of MLD and UCLD are compared, asymptotically as $p \to 0$, for the special class of quasiperfect linear codes. However, in general, the decoding rule which gives the minimum BER is difficult to deal with (see [12] for a clever reporting) and the improvement in the code's performance is not always significant, despite the increased complexity. Therefore it is of some interest to have an expression for the bit error probability for block codes over the binary symmetric channels when UCLD is applied.

Here we will show that the BER for block codes having a transitive group of automorphisms G can be obtained from the weight distribution by means of a differential linear transformation. It must be observed that several important classes of codes satisfy this requirement; some of the more common ones are cyclic codes, extended primitive binary BCH codes, uniformly packed linear codes, and shortened codes of cyclic codes which have a 2-transitive group of automorphisms; see [2, 5]. The condition of the automorphism group for the codes will allow a simple counting argument which will be used in the following to evaluate the bit error rate; otherwise, BER computation could require exhaustive analysis.

II. DEFINITIONS

This section is mainly devoted to recalling definitions as well as known results for easy reference. We just report, with few comments, the items useful in the computation of BER from the weight enumerator.

The leader weight enumerator is defined as the polynomial

$$L(X,Y) = \sum_{i=0}^{n} l_i X^i Y^{n-i},$$

where l_i is the number of coset leaders of weight i.

For every binary block code, linear or nonlinear, the bit error rate P_{symb} is defined, [2], as

$$P_{\text{symb}} = \frac{1}{kM} \sum_{i=1}^{k} \sum_{j=1}^{M} \text{Prob}\left\{x_i^{(j)} \neq \hat{x}_i \middle| X^{(j)} \text{ was sent}\right\}, \tag{1}$$

where M denotes the cardinality of the code, and the code vectors $X^{(j)}$, $j = 1, \ldots, M$, are equally likely. If attention is restricted to linear codes, then Equation (1) can be written in a simpler and useful form

$$P_{\text{symb}} = \frac{1}{k} \sum_{\mathbf{e}} f(\mathbf{e}) p^{\text{wt}(\mathbf{e})} (1-p)^{n-\text{wt}(\mathbf{e})}, \tag{2}$$

where $f(\mathbf{e})$ is the number of incorrect information bits after decoding with the assumption that the all zero word is transmitted, and the summation is extended all over the 2^n binary n-tuples. For computational purposes Equation (2) will be more conveniently rewritten as follows:

$$P_{\text{symb}} = \sum_{i=0}^{n} B_i p^i (1-p)^{n-i} = \sum_{i=0}^{n} E_i p^i, \tag{3}$$

where

$$B_i = \frac{1}{k} \sum_{\text{wt}(\mathbf{e})=i} f(\mathbf{e})$$

and $\Sigma_{\text{wt}(\mathbf{e})=i}$ sums $f(\mathbf{e})$ over all error patterns of Hamming weight i.

Now let $B(X, Y)$ denote the generating polynomial of the B_i's, that is,

$$B(X, Y) = \sum_{i=0}^{n} B_i X^i Y^{n-i}. \tag{4}$$

Thus it is seen that

$$P_{\text{symb}} = B(p, 1-p). \tag{5}$$

This last expression suggests the introduction of an operator L_1 which replaces X and Y respectively with p and $1-p$; it is immediately verified that L_1 is linear.

By using L_1, the expression (5) can be rewritten as

$$P_{\text{symb}} = L_1[B(X, Y)].$$

It is known that the word error probability P_w may be obtained from

$L(X,Y)$, whereas the probability P_d of detecting some errors at the receiver front end may be computed from $W(X,Y)$. In both cases using L_1 we respectively have

$$P_w = 1 - L_1[L(X,Y)]$$

and

$$P_d = 1 - L_1[W(X,Y)].$$

In the next section we will introduce a linear operator L_2 that enables us to mechanically derive $B(X,Y)$ from the weight enumerator $W(X,Y)$. We will show that the operator L_2 admits an explicit representation as antisymmetric homogeneous differential operator in the algebra $Z[[X,Y]]$.

III. COUNTING

As described in the previous section, the derivation of $B(X,Y)$ for systematic linear block codes is based on the evaluation of

$$B_j = \frac{1}{k} \sum_{\mathrm{wt}(\mathbf{e}) = j} f(\mathbf{e}),$$

where $f(\mathbf{e})$ is the number of 1's in information positions after the decoding when the received word is $\mathbf{e}$ and the all zero word is transmitted.

The evaluation of B_j is essentially a counting problem. Thus to this purpose we introduce the set of matrices $T^{(j)}$ whose rows are all the $\binom{n}{j}$ binary words of weight j, $0 \leqslant j \leqslant n$. In each matrix $T^{(j)}$ the rows are ordered so as to define a partition in $s \leqslant 2t+2$ submatrices $Q_1, Q_2, \ldots, Q_s$, where each submatrix is made up of words coming from all the codewords of a fixed given weight and modified by correctable suitable error patterns. The rows of each Q_i, $1 \leqslant i \leqslant s-1$, are vectors of the form

$$Z = C + E,$$

where C is a codeword and E is a coset leader, i.e. a correctable error pattern. Let $j + h_2 - h_1$ be the weight of C; let $h_1 + h_2$ be the weight of E, where h_1 denotes the number of 1's falling in the 0's positions of C, and h_2 denotes the number of 1's falling in 1's positions of C, so that the weight of Z

will be $\mathrm{wt}(C)+h_1-h_2=j$; and finally let h_2-h_1 be fixed and different from zero. The rows of the last submatrix Q_s are either words coming from codewords of weight j, or words that according to UCLD are decoded without changes in information positions.

For later use it is useful to know the number of rows in each matrix Q_i. Set $b=\lfloor (t-h)/2 \rfloor$; if $h_1 > h_2$, then the number N_{j-h} of rows of the submatrix associated to codewords of weight $j-h$, is given by

$$N_{j-h}=A_{j-h}\sum_{h_2=0}^{b}\binom{n-j+h}{h_2+h}\binom{j-h}{h_2}. \tag{6}$$

Similarly, if $h_1 < h_2$, the number N_{j+h} of rows of the submatrix associated to codewords of weight $j+h$ is given by

$$N_{j+h}=A_{j+h}\sum_{h_1=0}^{b}\binom{n-j-h}{h_1}\binom{j+h}{h_1+h}. \tag{7}$$

The number of rows of Q_s is given by the difference

$$N_s=\binom{n}{j}-\sum_{n=1}^{t}\left[N_{j-h}+N_{j+h}\right].$$

Every matrix defined above has the same number of 1's in all its columns. This statement can be proved as follows. In $T^{(j)}$ all the columns have the same number of 1's, since the interchange of two columns leaves the set of rows invariant. A similar argument also applies to any submatrix Q_i of $T^{(j)}$. Let us consider first the matrices Q_i, $1 \leqslant i \leqslant s-1$, where, as said before, every row has the form $C+E$, in which C is a codeword of weight $j-h_2+h_1$ and E is a coset leader of weight h_1+h_2. Let $P(g)$ denote a permutation matrix associated to the permutation g of G, and let $ZP(g)$ be the vector that results from applying the permutation g to the n components of Z. We have $(C+E)P(g)=CP(g)+EP(g)$, where $CP(g)$ is still a codeword and $EP(g)$ is a vector of weight not greater than t that modifies $CP(g)$ in the same way that E modifies C. Note that if E is a coset leader, of weight not greater than t, then $EP(g)$ is a coset leader too. Therefore $(C+E)P(g)$ is just transformed into another row of Q_i. This implies the existence of a permutation matrix $R(g)$ such that

$$R(g)Q_iP(g)=Q_i \qquad \text{for every } g \in G \text{ and for every } i \leqslant s-1,$$

and we can conclude that every matrix Q_i, $1 \leqslant i \leqslant s-1$, has the same number of 1's in every column. Finally, also the matrix Q_s has the same number of 1's in every column, because this number is the difference of equal numbers.

Now we are ready to estimate B_j. Owing to previous results, the contribution coming from matrices Q_i, $1 \leqslant i \leqslant s-1$, is obtained by summing the numbers

$$N_{j-h}(j-h)+N_{j+h}(j+h)$$

for h from 1 to t, then multiplying by the ratio k/n to take account of the k information columns, and finally dividing by k:

$$B_{1j}=\frac{1}{n}\sum_{h=1}^{t}\left[N_{j-h}(j-h)+N_{j+h}(j+h)\right].$$

The contribution coming from the last matrix Q_s is simply obtained multiplying the number N_s by j/n:

$$B_{0j}=\frac{j}{n}\left\{\binom{n}{j}-\sum_{h=1}^{t}\left[N_{j-h}+N_{j+h}\right]\right\}.$$

In conclusion we have

$$\begin{aligned} B_j &= B_{0j}+B_{1j} \\ &= \frac{1}{n}\left\{j\binom{n}{j}+\sum_{h=1}^{t}h\left[N_{j+h}-N_{j-h}\right]\right\}. \end{aligned} \tag{8}$$

Multiplying this expression by X^jY^{n-j} and summing over j, we get

$$\begin{aligned} B(X,Y) &= \frac{1}{n}\sum_{j=0}^{n}X^jY^{n-j}\left\{j\binom{n}{j}+\sum_{h=1}^{t}h\left[N_{j+h}-N_{j-h}\right]\right\} \\ &= X(X+Y)^{n-1}+\frac{1}{n}\left\{\sum_{j=0}^{n}X^jY^{n-j}\sum_{h=1}^{t}h\left[N_{j+h}-N_{j-h}\right]\right\}. \end{aligned} \tag{9}$$

The final expression for $B(X,Y)$ is obtained after some algebraic manipulations outlined in the following steps:

Step 1. Interchange summations in Equation (9) and use Equations (6) and (7):

$$X(X+Y)^{n-1}+\frac{1}{n}\sum_{h=1}^{t}\sum_{h_2=0}^{b}\sum_{j=0}^{n} hX^jY^{n-j}$$

$$\times\left[A_{j+h}\binom{n-j-h}{h_2}\binom{j+h}{h_2+h}-A_{j-h}\binom{n-j+h}{h_2+h}\binom{j-h}{h_2}\right].$$

Step 2. Perform the substitution $h=u-2h_2$, define $v=\lfloor u/2\rfloor$, separate the two summations on j, and in them respectively substitute $j=i+2h_2-u$ and $j=i-2h_2+u$:

$$X(X+Y)^{n-1}+\frac{1}{n}\sum_{u=1}^{t}\sum_{h_2=0}^{v}(u-2h_2)$$

$$\times\left\{\sum_{i=u-2h_2}^{n} X^{i+2h_2-u}Y^{n-i-2h_2+u}A_i\binom{n-i}{h_2}\binom{i}{u-h_2}\right.$$

$$\left.-\sum_{i=0}^{n-u+2h_2} X^{i-2h_2+u}Y^{n-i+2h_2-u}A_i\binom{n-i}{u-h_2}\binom{i}{h_2}\right\}.$$

Step 3. Rewrite the expression of step 2 to enhance the action of differential operators as follows:

$$X(X+Y)^{n-1}+\frac{1}{n}\sum_{u=1}^{t}\sum_{h_2=0}^{v}(u-2h_2)$$

$$\times\left\{X^{h_2}Y^{u-h_2}\sum_{i=u-2h_2}^{n} X^{i+h_2-u}Y^{n-i-h_2}A_i\binom{n-i}{h_2}\binom{i}{u-h_2}\right.$$

$$\left.-Y^{h_2}X^{u-h_2}\sum_{i=0}^{n-u+2h_2} X^{i-h_2}Y^{n-i+h_2-u}A_i\binom{n-i}{u-h_2}\binom{i}{h_2}\right\}.$$

Step 4. Finally find

$$B(X,Y) = X(X+Y)^{n-1} + [L_2(t)\{W(X,Y)\}], \tag{10}$$

where the linear differential antisymmetric homogeneous operator $L_2(t)$, associated with t-error correcting codes, has been introduced:

$$L_2(t) = \frac{1}{n} \sum_{u=1}^{t} D_u \tag{11}$$

with

$$D_u = \frac{1}{u!} \sum_{h_2=0}^{v} (u - 2h_2) \binom{u}{h_2} \times \left(Y^{h_2} X^{u-h_2} \frac{\partial^u}{\partial Y^{u-h_2} \partial X^{h_2}} - X^{h_2} Y^{u-h_2} \frac{\partial^u}{\partial X^{u-h_2} \partial Y^{h_2}} \right)$$

In particular the first three expressions for $L_2(t)$, corresponding to minimum distances from 3 up to 8, are

$$L_2(1) = \frac{1}{n} \left(Y \frac{\partial}{\partial X} - X \frac{\partial}{\partial Y} \right),$$

$$L_2(2) = \frac{1}{n} \left[\left(Y \frac{\partial}{\partial X} - X \frac{\partial}{\partial Y} \right) + \left(Y^2 \frac{\partial^2}{\partial X^2} - X^2 \frac{\partial^2}{\partial Y^2} \right) \right],$$

$$L_2(3) = \frac{1}{n} \left[\left(Y \frac{\partial}{\partial X} - X \frac{\partial}{\partial Y} \right) + \left(Y^2 \frac{\partial^2}{\partial X^2} - X^2 \frac{\partial^2}{\partial Y^2} \right) + \frac{1}{2} \left(Y^3 \frac{\partial^3}{\partial X^3} - X^3 \frac{\partial^3}{\partial Y^3} \right) + \frac{1}{2} \left(XY^2 \frac{\partial^3}{\partial X^2 \partial Y} - YX^2 \frac{\partial^3}{\partial Y^2 \partial X} \right) \right].$$

It is interesting to observe that $L_2(t)$ is the sum of t differential operators,

where each operator originates from the correction of all error patterns of a given weight i, for every i in the range between 1 and t.

IV. CONCLUSIONS

In this paper we have introduced the operator $L_2(t)$ that enables us to obtain the bit error probability from the weight enumerator. The existence of $L_2(t)$ is seen to follow from the uniqueness of coset leaders and the transitivity of the symmetry permutation group. These requirements are not mandatory; they are sufficient conditions, and so far easy counterexamples show that the transformations may either work or not with codes without a transitive group of symmetry. Hence necessary and sufficient conditions on the code for Equation (10) to work will be welcome. Moreover, several other interesting questions still remain open, for example a combinatorial interpretation of the surprising symmetry of the operator $L_2(t)$.

There exists a nice commutation property between D_1 and M_c, i.e.

$$\left(Y\frac{\partial}{\partial X} - X\frac{\partial}{\partial Y}\right)M_c = -M_c\left(Y\frac{\partial}{\partial X} - X\frac{\partial}{\partial Y}\right),$$

which can be verified by straightforward computation. It can be used to compute the BER of one-error-correcting codes from the weight enumerators of their dual codes. For example, for Hamming codes, the BER, known from [10], can be simply obtained given the weight enumerator of their dual codes:

$$W(X,Y)^{\perp} = X^n + nX^{(n-1)/2}Y^{(n+1)/2}.$$

The following steps are straightforward. Apply $L_2(1)$ to $W(X,Y)^{\perp}$:

$$L_2(1)W(X,Y)^{\perp} = \frac{1}{n}\left[nYX^{n-1} + \frac{n(n-1)}{2}X^{(n-3)/2}Y^{(n+3)/2} - \frac{n(n+1)}{2}X^{(n+1)/2}Y^{(n-1)/2}\right]. \quad (12)$$

Apply M_c to Equation (12):

$$Q(X,Y) = M_c L_2(1) W(X,Y)^{1}$$

$$= \frac{1}{n}\Bigg[n(X-Y)(X+Y)^{n-1} + \frac{n(n-1)}{2}(X+Y)^{(n-3)/2}(X-Y)^{(n+3)/2} - \frac{n(n+1)}{2}(X+Y)^{(n+1)/2}(X-Y)^{(n-1)/2}\Bigg].$$

TABLE 1
COEFFICIENTS FOR BER COMPUTATION OF SOME KNOWN (n, k, d) CODES

	E_i						
i	(7, 4, 3)	(8, 4, 4)	(12, 6, 4)	(15, 10, 4)	(15, 5, 7)	(22, 11, 6)	(24, 12, 8)
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	9	7	11	14	0	0	0
3	− 26	− 14	− 50	− 63	0	210	0
4	30	0	120	28	469	− 2555	1771
5	− 12	28	− 144	882	− 3948	16086	− 21252
6	0	− 28	0	− 4508	15918	− 64848	102718
7	0	8	288	12496	− 39780	175120	− 118404
8		0	− 480	− 23456	67410	− 289140	− 1445136
9			400	31920	− 79968	87640	10852688
10			− 176	− 32032	66220	1120448	− 43878296
11			32	23296	− 36792	− 4106592	126252336
12			0	− 11648	12376	8918000	− 281248968
13				3584	1904	− 14166880	504040768
14				− 512	0	17429760	− 740844720
15				0	0	− 16903168	900348064
16						12909120	− 904655136
17						− 7647360	746030208
18						3404800	− 497466816
19						− 1075200	261824640
20						215040	− 104729856
21						− 20480	29922816
22						0	− 5440512
23							473088
24							0

Use Equation (10):

$$B(Y) = X(X+Y)^{n-1} - \frac{1}{n+1}Q(X,Y).$$

Apply L_1:

$$P_{symb} = \frac{1}{2^m}\left\{1+(2^m-2)p-(1-2p)^{2^{m-1}-1}\left[1+2(2^m-2)(p-p^2)\right]\right\}.$$

Finally using Equation (10), the BER of several codes has been derived, in particular the known expressions of all perfect codes have been checked. Numerical results obtained using the program MUMATH are summarized in Table 1, where are reported the coefficients E_i to be used in equation (3) for evaluating the BER.

I would like to acknowledge the kind suggestions of the editor and referee, which have improved the formulation of the paper.

REFERENCES

[1] D. Slepian, A Class of binary signaling alphabets, *Bell System Tech. J.*, 35:203–234 (Jan. 1956).

[2] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error Correcting Codes*, North-Holland, Amsterdam, 1977.

[3] F. J. MacWilliams, A Theorem on the distribution of weights in a systematic code, *Bell System Tech. J.*, 42:79–94 (1962).

[4] J. H. van Lint, *Introduction to Coding Theory*, Springer, New York, 1982.

[5] E. Berlekamp, *Algebraic Coding Theory*, McGraw-Hill, New York, 1968.

[6] Shu Lin and D. J. Costello, *Error Control Coding: Fundamentals and Applications*, Prentice-Hall, Englewood Cliffs, N.J., 1983.

[7] M. Elia, Symbol error rate of binary block codes, in *Transactions of the Ninth Prague Conference on Information Theory, Statistical Decision Functions, and Random Processes*, Prague, June 1982, pp. 223–227.

[8] M. Elia and G. Prati, On the complete decoding of binary linear codes, *IEEE Trans. Inform. Theory*, IT-31(4):518–520 (July 1985).

[9] G. Rota, *Studies in Combinatorics*, Math. Assoc. Amer., 1978.

[10] J. H. van Lint, *Coding Theory*, Lectures Notes, Springer, New York, 1972.

[11] W. W. Peterson and E. J. Weldon, *Error-Correcting Codes*, MIT Press, Cambridge, Mass., 1981.

[12] L. A. Dunning, Encoding and decoding for minimization of message symbol error rates in linear block codes, *IEEE Trans. Inform. Theory*, IT-33(1):91–104 (January 1987).

[13] N. J. A. Sloane, *A Short Course on Error-Correcting Codes*, CISM course 188, Springer, Wien, 1975.

[14] J. F. Blake and R. C. Mullin, *The Mathematical Theory of Coding*, Academic, New York, 1975.

Received 13 January 1987; revised 23 June 1987